

Evaluation of the behavioral and affective outcomes of novice teachers working in professional development schools versus non-professional development schools

This is the accepted version, accepted for publication in Studies in Educational Evaluation, see [10.1016/j.stueduc.2017.10.006](https://doi.org/10.1016/j.stueduc.2017.10.006) for the published version

Michelle Helms-Lorenz*, Wim van de Grift¹, Esther Canrinus², Ridwan Maulana, Klaas van Veen
Department of Teacher Education, University of Groningen, The Netherlands

* Corresponding author at: Department of Teacher Education, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands.

1 Present address: Van Lyndenlaan 10a, 3768 MG Soestduinen, The Netherlands.

2 Present address: University of Agder, Department of Education, Kristiansand, Norway.

Acknowledgements

With gratitude we acknowledge the contribution of the schools, students, their teachers, as well as the assistance of dr. Bert Slof and dr. Marieke Alferink to this project. This study was funded by The Netherlands Organisation for Scientific Research (NWO project number 411-09-802).

Abstract

The purpose of this study was to evaluate the outcomes of teacher preparation based in Professional Development Schools (PDSs) in secondary education in The Netherlands over a period of three years. Approximately 150 teachers in non-PDSs were compared with 50 teachers in PDSs with declining response rates. Classroom observation ratings and student perceptions of PDS teachers in the second year were higher compared to non-PDS teachers. PDS teachers reported higher levels of teacher efficacy in each year. PDS-based teachers evaluated their teacher preparation and their learning opportunities significantly more positively compared to non-PDS-based teachers. Even though the sample size is not large and the design of the study does not warrant causality claims, the findings point towards increased transition smoothness and improved adhesiveness between teacher education preparation and the requirements on the job. Future investments in the co-operation between schools and teacher education institutes seem warranted.

Keywords

Teacher-education, program evaluation, school-based evaluation, secondary school teachers, professional development school

Introduction

This study sets out to investigate the quality of teaching skills and well-being during the transition from being a student teacher to the first years on the job. Beginning teachers and schools have complained about the lack of alignment between the teacher education curriculum on the one hand and the teaching skills and knowledge needed in the schools and classrooms on the other hand causing transition shock, “painful beginnings” and high novice attrition rates (e.g., Corcoran, 1981; Huberman, 1989; Ingersoll & Strong, 2011). Den Brok, Wubbels and van Tartwijk (2017) explored self-reported attrition causes and revealed that causes in The Netherlands were similar to those revealed in international studies: being related to the characteristics of the teaching job and the local school context in which beginning teachers find themselves and personal reasons and, in many cases, it is a combination of factors rather than a single factor that causes attrition. Intensified collaboration between schools and teacher education institutes has become the focus for improving the alignment between theory and practice before professional certification. Induction arrangements have been developed to improve the support after certification (e.g., Ingersoll et al., 2011). Co-operation

between schools and teacher education during and after teacher preparation brings together the expertise needed to enhance the development of teaching skills, and prevents stagnation of development. This way the support is constructed on both sides by both parties, i.e. the school and the education institute. In an experimental setting, Helms-Lorenz, van de Grift and Maulana (2015) revealed that after certification, attrition rates are reduced and the development of teaching skills is accelerated for teachers who teach in schools that work together with teacher education institutes to develop comprehensive teacher induction programs for their beginning teachers. Our study focusses on the longitudinal influence of intensified collaboration between schools and teacher education institutes on observed teaching skills and self-reported well-being of novices *before* qualification.

The importance of teacher education in general

Recent American and Dutch research shows the importance of teacher education programs. The better the teachers are educated in a formal teacher education program, the more effective they are, especially compared with those who did not receive any specific training (Boyd et al., 2009; Darling-Hammond & Youngs, 2002, Helms-Lorenz et al., 2015). Consistently, The Dutch Inspection of Education (2011) concluded that teachers who are not certified and not educated in a formal teacher education program pose an increasing problem for the quality of education. According to Dronkers (2010), who analysed the PISA-results, Dutch pupils would score substantially higher if all schools would have certified teachers trained in formal teacher education program. Helms-Lorenz et al. (2015) revealed that certification and level of teaching skills at the beginning of the career were important factors explaining teacher retention.

Professional Development Schools (PDS)

The general aims of partnerships between schools and teacher education institutes in the United States, called Professional Development Schools (PDSs), are to develop school practice; including the practice of all stakeholders, e.g. pre-service, beginning and experienced teachers (Nath, Guadarrama, & Ramsy, 2011), and that of school leaders. This is achieved through collaboration between novice teachers and experienced teachers and collaboration between school-based educators and university-based faculty (Darling-Hammond, 2005). The second aim of PDSs is to bridge the gap between the professional preparation and the real world of school reform, according to the National

Council for the Accreditation of Teacher Education (2001). This gap between the ideas pre-service teachers have about teaching (developed from prior experience and from teacher education) and their first experiences as a teacher in a school, generally causes friction, which may result in a decline in motivation for the profession (Flores, 2001). Lower levels of motivation, in turn, may result in teachers leaving the profession as motivation has been shown to be related to teachers' professional commitment (Canrinus, Helms-Lorenz, Beijaard, Buitink, & Hofman, 2012) and job stress (Davis & Wilson, 2000). Bridging this transition gap is one of the goals of PDS.

Teacher education in the Netherlands has a long history of a dominant practical component of school engagement in the curriculum. The aims of PDSs in The Netherlands are similar to the Anglo-Saxon aims: 1) to bridge the gap between the job requirements and the theoretical curriculum requirements (to reduce the transition shock from student teacher to becoming a teacher), and 2) to reduce the theory/practice gap experienced by novices. Schools prioritize the organization of pupil learning whilst teacher educators prioritize pre-service teacher learning. These aims can lead to contradicting and confusing practices for the novice teacher causing tensions in the pre-service teacher's learning process; e.g., by the school giving the novice too much responsibility (to treat the novice as an equal). Pre-service teachers struggle to teach and are pre-occupied with surviving the day, whilst experiencing the theoretical underpinnings as an extra burden if these are not connected to their reality.

The Dutch PDS is composed of one or more schools and one or more teacher education institutes forming a partnership and taking shared responsibility for educating pre-service teachers. In the Dutch context the shared responsibility, between the school and the teacher education institute is considered to be the best basis for teacher education. This partnership allows for handling the tensions between practical and theoretical insights. It entails structural appointments concerning job openings for pre-service teachers, organizing school-based teacher support staff, meetings discussing the teacher education curriculum and assessment procedures. Recent Dutch studies reveal that there is room for improving the quality of PDSs in the Netherlands, as the partnerships are capable of making structural appointments but are less capable of developing the curriculum and educating pre-service teachers in the same jargon and with the same goals and methods (e.g., Timmermans, 2012; Kroeze, 2014).

Effective elements of PDS

Kruger, Davies, and Eckersley (2009) studied seven effective and sustainable university-school partnerships in Australia to grasp which characteristics make these partnerships effective and

sustainable. They concluded that an effective PDS “[...] encourage(s) each stakeholder (pre-service teachers, teachers, teacher educators) to contribute personal and professional resources, in the form of passion, commitment and professional understanding and expertise.” (p. 10). Effective partnerships are characterized by 1) a focus on learning for all stakeholders, 2) all stakeholders taking on altered relationship practices, i.e., conversations aimed at improving the learning of school students, and 3) a focus on forming new relationships between partners (Kruger et al., 2009). Mantle-Bromley (2001) described five features of PDS-based teacher preparation that were thought to lead to more well-prepared teachers: (a) field experiences are spread throughout the teacher preparation program, (b) theory and practice are closely connected, (c) pre-service teachers are supervised by both the school and the institute personnel throughout the program, (d) a strong support system is provided, (e) the program is collaboratively designed and implemented.

Besides the ambitious aims of PDSs, possible drawbacks should be considered. One of the dangers Hargreaves (2000) anticipated is that limiting the period and level of certification for initial teacher education, or shifting the burden of “professional” preparation from higher education to the already overburdened world of schools (Barton, Barrett, Whitty, Miles, & Furlong, 1994), could reduce the opportunity for critical reflection. “The effect of this all can be to return teaching to an amateur, de-professionalized, almost premodern craft, where existing skills and knowledge are passed on practically from expert to novice” (p. 168). He advised in this respect that teachers should...

“[...] value and defend their entitlement *to* and their education in *a rigorous knowledge base that undergirds their professionalism*. [...] Engaging with such a knowledge base uplifts teachers from the pre-professional prejudice that only practice makes perfect. [...] It (the knowledge base) must also be treated as providing forms of educational understanding, and ways of assessing and filtering educational research, rather than falsely deifying and uncritically applying a body of incontrovertible scientific ‘fact’ on effective teaching, learning, management, and change strategies.” (p. 170). ...“Attempts to dismantle or destabilize the university base of teacher education should be interpreted and resisted as constituting not only an assault on teacher education, but on the professional status of teaching itself” (p. 171).

Given the conflicting expectations of PDS, little quantitative evidence is available supporting the claim that PDS programs provide superior learning environments for teacher preparation. Studies investigating whether the aims of PDS are achieved, i.e., improving teaching practice and reducing the transition shock between education and work, are scarce (Darling-Hammond, 2005). To date, no studies were found evaluating the PDS aims longitudinally and in concert, but rather studies

evaluating single aims in cross-sectional designs. These studies are presented in the following section.

The influence of PDS on the quality of teaching skills

A study by Castle, Fox, and Souder (2006) addressed the assumption that PDS educated teachers had more experience in teaching and thus would be better on the No Child Left Behind teacher quality indicators (i.e., planning, instruction, classroom management, assessment, professionalism and reflection). Based on student teaching evaluation forms and portfolio assessments, Castle et al. (2006) found differences between PDS teachers and non-PDS teachers. PDS teachers were rated significantly higher on their teaching regarding instruction, classroom management, and assessment. In their portfolios, the PDS teachers showed higher levels of sophistication when discussing the teacher standards.

In a number of small scale studies ($N=14$ to 79), PDS graduates have shown superior basic skills such as productive instruction (Sharpe, Lounsbery, Golden, & Deibler, 1999) and instructional strategies (Houston, Hollis, Clay, Ligons, & Roff, 1999). Wait and Warren (2001) and Wait (2000) found higher ratings of classroom-management skills, instructional skills, facilitation, and feedback in a three year follow-up study for PDS vs. campus-based elementary education graduates.

Ridley, Hurwitz, Davis Hackett, and Knutson Miller (2005) conducted a small-scale cross-sectional two year study comparing the lesson planning, teaching effectiveness, post-lesson reflectivity, and content retention of professional teaching knowledge for elementary teachers prepared at a PDS or campus-based program. They studied two phases: student teaching and first year teaching. Although they found that the scores of PDS-prepared student teachers consistently trended higher than the campus-prepared cohort, no statistically significant differences were found over 2 years. However, during the 1st year of teaching, PDS-prepared teachers scored significantly higher than campus-prepared teachers on teaching effectiveness.

These findings lead to our hypothesis that PDS-based pre-service teachers' teaching skill ratings will be higher during the first year of teaching in The Netherlands. The central assumption that PDSs are better learning environments for pre-service teachers compared to non-PDSs is scrutinized in this study.

The influence of PDS on teaching efficacy

Sandholtz and Dadlez (200) observed a very slight, yet significant, gain in the self-reported teaching efficacy in their study of 73 pre-service teachers after one year of studying in a PDS program. Several

years later, Tournaki (2009) and colleagues, although not specifically including a PDS-program, did not observe significant differences between a traditional certification program and two alternative certification programs in pre-service teachers' sense of teacher efficacy. Yet, Swars and Dooley (2010) revealed a significant increase in personal science teaching efficacy of 21 elementary science teachers in a PDS-based science methods course. This leads to our hypothesis that in the Netherlands, PDS-based pre-service teachers' self-efficacy will be higher compared to the self-efficacy of non-PDS-based pre-service teachers.

The influence of PDS on teacher stress

Numerous studies have shown that teacher stress is related to physical illnesses, mental illnesses and reduced teacher commitment and effectiveness. Stress and well-being at work are related to teaching practices (Harmsen, Maulana, Helms-Lorenz, & van Veen, under review). Teaching practices have been shown to relate positively to job retention (Helms-Lorenz et al., 2015). No previous research was found comparing stress levels of teachers in PDSs vs non-PDSs. Therefore we will explore this relationship without a hypothesis.

The influence of PDS on commitment to teach

Results of a study by Latham and Vogt (2007) revealed that while controlling for several individual variables, PDSs significantly and positively affected how long teachers stayed in education. Although these are positive findings, Latham and Vogt were unable to control for self-selection based on motivation or commitment, as their study included teachers who voluntarily participated in a PDS-trajectory for obtaining a teaching qualification. They did present small scale, preliminary findings indicating that commitment and motivation did not differ across groups (Latham & Vogt, 2007). Sandholtz et al. (2000) observed, in an early study on the effects of PDSs, a substantial decline in 73 pre-service teachers' commitment after one year a PDS program. Then again, they did not compare these findings with data of students from traditional programs. As such, this decrease in commitment cannot completely be linked to the PDS program. In addition, Mantle-Bromley, Gould, McWhorter, and Whaley (2000), on the other hand, did not find significant differences in pre-service teachers' intent to stay in their current job, number of teaching jobs held, and their intent to search for new positions across three certification programs, amongst which was one PDS program. Thus, the question remains whether the pre-service teachers who obtained their teaching certification through a PDS program stay in the teaching profession because the PDS program influenced their commitment to their occupation. We will explore this relationship without a hypothesis.

The influence of PDS on job satisfaction

Little is known about the relationship between teachers' job satisfaction and being educated to become a teacher through a PDS program. What has been shown is that a relationship exists between the aforementioned commitment components and job satisfaction (Cooper-Hakim & Viswesvaran, 2005), with job satisfaction influencing occupational commitment (Carrinus et al., 2012; Landsman, 2001). In a study by Paese (2003), job satisfaction was included as a subscale of the measurement of task based stress. Using a longitudinal design, Paese did not observe significant pre- to post-test changes in the 48 pre-service teachers' job satisfaction. This was the case for the 24 pre-service teachers following a semi-PDS program as well as for the other 24 pre-service teachers following a full PDS program. The differences between these two groups of teachers were not significant either. Unfortunately, this is only one study, with a restricted sample size and without comparison to a sample of pre-service teachers educated through traditional teacher education. As such, it remains unclear if, and if so to what extent, there is a relation between pre-service teachers' job satisfaction and being educated in a PDS. Therefore we will explore this relationship without a hypothesis.

The influence of PDS on feeling well-prepared for the job

In general, PDS graduates in the United States feel well-prepared, confident and are less likely to experience a "culture shock" (Blocker & Mantle-Bromley, 1997; Book, 1996; Patterson, 2000; Ross, 2001; Sandholtz et al., 2000; Thompson & Ross, 2000; Trachtman, 1996; Tusin, 1995; Walling & Lewis, 2000; Yerian & Grossman, 1997). It remains unclear whether this is the case in the Dutch context. Therefore we will explore this relationship without a hypothesis.

Research aims

Because of a gap in the literature regarding the longitudinal effect of PDSs on the quality of teaching skills of pre-service teachers in general and in the Netherlands in particular, this research aims to enlarge the body of evidence regarding the added value of co-operation between teacher education institutes and schools in secondary education by incorporating behavioural as well as psychological aspects.

This study aims to evaluate teacher preparation routes in the Netherlands. In particular, a comparison is made between the teaching behaviour and psychological indicators of transition smoothness of PDS- and non-PDS-based pre-service teachers over a period of three years.

As the evidence found in the literature with regard to the development of stress, job-commitment and job satisfaction is inconclusive, hypotheses cannot be formulated and therefore, as mentioned above, explorative research questions are formulated. With regard to teaching skills and self-efficacy, specific hypotheses seem warranted based on previous research findings.

Hypotheses

1. The level of teaching skills of pre-service teachers in PDSs is higher than that of pre-service teachers in non-PDSs.
2. The level of self-efficacy of pre-service teachers in PDSs is higher than that of pre-service teachers in non-PDSs.

Explorative research questions

3. Are the reported stress causes and responses of PDS-based pre-service teachers the same over a period of three years compared to that of non-PDS-based pre-service teachers?
4. Is the job-commitment of PDS-based pre-service teachers the same over a period of three years compared to that of non-PDS-based pre-service teachers?
5. Is the job satisfaction of PDS-based pre-service teachers the same over a period of three years compared to that of non-PDS-based pre-service teachers?
6. Are PDS-based pre-service teachers equally satisfied with the way they were prepared for teaching?

Method

Participants and data collection procedures

In this study we compare secondary teachers who were educated in different institutes spread over the Netherlands. The variable of interest is the school of internship: PDS or not. The teacher education program is equal for the students, except for the internship school's responsibility and involvement in educating the pre-service teacher. This study was part of a national, longitudinal study on 'The effects of educating teachers at school', funded by the Netherlands Organization for Scientific Research (NWO project number 411-09-802). School leaders and coordinating 'school educators' of all secondary schools (N=650) in the Netherlands were approached in 2009 by letter and telephone to participate in the study. After the schools confirmed to participate, they provided contact information of pre-service teachers in their school or institution. A total of 657 pre-service teachers (approximately 22% of the pre-service teachers in the Netherlands) with their field placement in 87 participating schools (approximately 13% of the Dutch secondary schools) were contacted by email, and informed about the study goals, procedures, and data collection. The response rate differed per measurement scale and time point; therefore, samples sizes are mentioned above each table in the results section. Participation of schools and pre-service teachers was voluntary. The participating schools were grouped into the PDS group when they were qualified as a PDS by the NVAO (see NVAO, 2009, for qualification criteria).

Experienced teachers and school educators were selected to participate in a four hour observation training. This training explains the observation instrument and scoring guidelines, and observers practice using two recorded lessons. After the first observed recorded lesson, the percentage of consensus between the observers was calculated. Then, differences in scoring between observers was discussed, as well as the agreement between the participant group and a previously established norm group. A consensus of 70% or higher on the second observed lesson is aimed for.. In the present study, this criterion was met.

Data collection was longitudinal, throughout three consecutive academic years (starting in 2010). In the first data-collection year all the participants were in the last phase of their teacher preparation. In the second year they were very recently certified and in the last year they were in their second year on the job as a certified teacher. Three waves of data were collected; the first wave starting in September 2010 (2010-2013), the second wave in September 2011 (2011-2014) and the third wave in September 2012 (2012-2015). Data were collected at fixed time points; the questionnaires were

surveyed across the year with monthly intercessions. Peer observations and student questionnaires were collected in January and February.

Professional development schools in the context of the Netherlands

Since the start of this century, teacher education in the Netherlands has shifted, in line with other countries, in the direction of connecting teacher education with schools. In 2001, the Dutch Education Council advised to grant schools more influence within the educational infrastructure (Onderwijsraad, 2001). First, suggestions towards a more collaborative perspective on educating pre-service and in-service teachers were made (Onderwijsraad, 2001). This implied more collaboration between schools and universities and/or teacher training institutions. The Dutch ministry of education, culture and science responded to this recommendation with a plan. In this plan, the school was perceived as the central player in the field of education. The school was to search for and request support and guidance from external organizations regarding their Human Resource Development-policy. This would ensure further development of these school organizations (Dutch Ministry of Education, Culture and Science, 2002).

Between 2002 and 2006, the government funded a small number of pilot schools to develop an infrastructure in which workplace learning or school-based learning could be implemented (Educational Inspectorate, 2007). From 2005 onwards, studies were performed to investigate how the quality of these school-based teacher education agreements should be assessed and assured (CFI, 2005). This resulted in national criteria for professional development schools (Dutch Ministry of Education, Culture and Science, 2008) which the Dutch and Flemish accreditation organization (NVAO) used to develop an assessment framework to judge the quality of partnerships developed between schools and universities or teacher training institutions (NVAO, 2009).

Using this framework, the government officially qualified and funded the first 58 partnerships in primary, secondary, and vocational and adult education as PDSs in 2009. The present study sets out to investigate to what extent PDS schools and non-PDS schools contribute differentially to the level of their pre-service teachers' teaching quality and the transition to the teaching profession.

In the context of the Netherlands, the five features of PDS-based teacher preparation, described by Mantle-Bromley (2001), are upheld: (a) field experiences are spread throughout the teacher preparation program, (b) theory and practice are closely connected, (c) pre-service teachers are supervised by both the school and the institute personnel throughout the program, (d) a strong support system is provided, and (e) the program is collaboratively designed and implemented.

Concepts and instruments used to evaluate PSD in the Dutch context

Teaching quality: peer observation

Over the last two decades, evaluating teachers' teaching effectiveness has been a major focus in educational research. The main reason for this is that teachers' teaching effectiveness is an important factor for pupils' learning and outcomes (Creemers & Kyriakides, 2008; Hattie, 2012). Effective teaching behaviour is defined as behaviour associated with higher learning and outcomes in pupils, such as academic engagement, motivation and achievement (Maulana, Helms-Lorenz, & van de Grift, 2015).

Research has shown multiple factors to be related to teaching quality, related to the teacher, the learning environment, and interpersonal contact. These factors operate at different levels, e.g., the national/regional, school, or classroom level (Kyriakides, Creemers, & Antoniou, 2009). Although we realize that factors at different levels contribute, and affect one another, the focus of this study is on effectiveness factors at the teachers' level, by comparing teaching behaviour from PDS teachers to teaching behaviour from non-PDS teachers.

Based on educational effectiveness research, and inspired by the theory of Fuller (1969), van de Grift, van der Wal and Torenbeek (2011) proposed six domains of effective teaching behavior. These behaviors are restricted to behaviors that are observable during a lesson. The behavioral domains are labelled as follows; providing a safe and stimulating learning environment, efficient classroom management, clarity of instruction, activating learning, adaptive teaching, and teaching learning strategies. This observational instrument was developed and validated in primary education, (van de Grift, 2007), as well as in secondary education (van de Grift, Helms-Lorenz, & Maulana, 2014). The items composing the six behavioral domains meet the assumptions of the Rasch model, implying that that behavioral domains can be ordered in terms of complexity level from 'easy' to 'difficult', and that easy teaching behavior is conditional for more advanced or difficult teaching behavior

Quality of teaching was measured by a 32-item observation instrument developed by van de Grift, (2007). The instrument consists of 6 domains. Each domain contains between 3 and 7 items, for example 'adapts instruction to relevant differences between pupils', which are rated on a four point Likert scale: 1 = *generally weak*, 2 = *more weak than strong*, 3 = *more strong than weak* and 4 = *generally strong*. Previous research has shown evidence of scale validity (van de Grift et al., 2014). The items scores were converted to Rasch scale scores in terms of Warm's Theta coefficients (van de Grift et al., 2014).

Three observation items were used to measure student's psychological and behavioral academic engagement. An example is; students show that they are interested in learning'. Items were rated on a four point Likert scale, ranging from 1 (*generally weak*), to 4 = (*generally strong*). Previous studies showed that internal consistency is high (.89 in a sample of 264 pre-service teachers) and that student academic engagement has predictive validity for pre-service teachers' skillfulness (Maulana, Helms-Lorenz, & van de Grift, 2014).

Teaching quality: student perceptions

Student perceptions of teaching quality were measured by a student version of the observation instrument by van de Grift (2007), called the "My teacher..." questionnaire, which was adapted and psychometrically tested in a previous study by Maulana, Helms-Lorenz and van de Grift, (2014). This questionnaire consists of 41 items, divided over the 6 behavioral domains described above. For this longitudinal study, the 24 items of the questionnaire, as presented in Maulana et al. (2014), that fitted the Raschscale were used. Previous research has shown the evidence of scale validity (Maulana & Helms-Lorenz, 2016). The items scores were converted to Rasch scale scores in terms of Warm's Theta coefficients (Maulana et al., 2014).

Teacher efficacy

A psychological indicator of interest for measuring a smooth transition is the teacher's self-efficacy: teachers' confidence in their abilities to provide clear instructions, providing effective classroom management and engaging students in learning. More in general, teaching efficacy (Tschannen-Moran, 1998) is defined as the belief in one's capabilities to organize and execute courses of action required to successfully accomplish a specific teaching task in a particular context (based on the work of Bandura on general self-efficacy). Studies have shown that teacher efficacy is related to various outcomes such as student achievement and job satisfaction (Caprara, Barbaranelli, Steca, & Malone, 2006).

Perceived sense of teacher efficacy was measured using the Teacher's Sense of Efficacy Scale (TSES; Tschannen-Moran & Woolfolk Hoy, 2001). This questionnaire consists of 24 statements, divided into three domains, respectively, *efficacy for instruction* (8 items, e.g. ...), *efficacy for classroom management* (8 items, e.g. ...) and *efficacy for student engagement* (8 items, e.g. ...). In addition, statements regarding the teacher's sense of collective efficacy were proposed using 6 questions on *collective efficacy for instruction*, and 6 questions on *collective efficacy for classroom management*. Responses were coded using a five-point Likert scale, ranging from 1 (*nothing*) to 5 (*a great deal*). Statements were provided with the following instruction: 'Indicate the extent to which each of the

statements is applicable to you or to the class. Please provide an estimation in case you do not have much experience as a teacher.’ The reliability (internal consistency) of the three subscales was found to be high (between .89 and .94) in previous studies (Tschannen-Moran & Woolfolk Hoy, 2001; Duffin, French, & Patrick, 2012).

Teacher stress

Teacher stress is defined in various ways, mostly incorporating work-related factors such as task demands, lack of support from colleagues and the organisation, and personal factors, such as the individual’s ability to cope with tasks. Teacher stress occurs when teaching demands are high, and the (perceived) resources are low (Lambert, McCarthy, O’Donnell, & Wang, 2009). A distinction is made between stress causes and stress responses, i.e. the psychological reaction to stress, for example, the (lack of) pressure an individual perceives as a result of the job demands (van Veldhoven & Meijman, 1994).

The Monitor at Work questionnaire (van Veldhoven, de Jonge, Broersen, Kompier, & Meijman, 2002¹) was used to measure perceived stress causes and responses. The extensive, 207-item version of the questionnaire was used, comprising 27 subscales, organized in 6 higher-order subscales. Four subscales measure stress causes (high psychological task demands, a lack of learning opportunities, a lack of regulating possibilities, and poor social-organizational aspects), and two subscales measure stress responses (tension and job discontent). The psychometric quality of the questionnaire has been tested extensively, showing good reliability and validity (construct, concurrent, criterion, and predictive validity) (van Veldhoven et al., 1993; van Veldhoven, 1996).

Job satisfaction

Job satisfaction is defined as “An attitude based on an evaluation of relevant aspects of the work and work situation.” (Van der Ploeg & Scholte, 2003, p. 277). Job satisfaction was measured using the job satisfaction index (ASI) constructed by van der Ploeg and Scholte (2003). The instrument measures 5 job satisfaction constructs using Likert scales: the satisfaction with support on the job, satisfaction with the amount of job-autonomy, satisfaction with relationships at work, satisfaction with the tasks, and with fringe-benefits and salary. In various previous studies the scale reliabilities have been documented having Cronbach’s alpha’s above .80 (e.g., Van der Ploeg & Scholte, 2003 Canrinus, 2011; Evers, 2013).

¹ The license and copyright of this questionnaire was obtained from SKB, Entrada 501, 1114 AA Amsterdam

Job commitment

Job-commitment refers to “a psychological link between a person and his or her occupation that is based on an affective reaction to that occupation” (p. 800, Lee et al., 2000). Occupational commitment was measured using an instrument developed by Meyer, Allen, and Smith (1993). The instrument measures 3 constructs: affective commitment referring to the positive emotions towards the profession, normative commitment referring to the sense of obligation to remain in the occupation, and continuance commitment referring to the awareness of the costs associated with leaving the occupation (Meyer et al., 1993). In various previous studies the Cronbach alpha’s were reported to range between .73 and .80 (e.g., Meyer, Allen, & Smith, 1993; Canrinus, 2011).

Perceived quality of teacher education

Another psychological indicator of a smooth transition is the extent of satisfaction experienced by the pre-service teacher with the contribution of the teacher education to job preparation. Teacher education provides transition support in the curriculum by organizing meetings with students to discuss concerns and classroom and school related difficulties, including support from the teacher education institute as well as the school-educator.

To measure the perceived quality of teacher education the National Student Survey (Nationale Studenten Enquête) was used. All tertiary students in the Netherlands are annually approached with this instrument to monitor their satisfaction with their tertiary education program. The students rate the general quality of the program, the acquired skills, acquired scientific skills, acquired practical skills, the career preparation, quality of lecturers/educators, study information supply, quality of study facilities, satisfaction with the assessment procedures, satisfaction with the schedule, rating of the study load, the provided support, and the personal engagement with the institute. The NSE is organised by Studiekeuze123. This foundation is the result of an initiative of Vereniging Hogescholen (Higher Education Council), NRTTO (Dutch Council of Training and Education), VSNU (University Association) as well as the student organisations ISO and LSVb. Research is carried out by the Dutch research institute GfK.

Background variables

Background variables were assessed by a questionnaire to the teacher. Table 1 presents the background characteristics (gender, age, and type of education) of the participating teachers. No differences in background characteristics were found between PDS teachers, non-PDS teachers, and teachers with missing data on the type of school they were educated.

Table 1 Background characteristics of participating teachers

Teacher's date of birth	First 25%^b (1953-1980)	Second 25% (1981-1986)	Third 25% (1987 - 1988)	Fourth 25% (1989 - 1993)	Difference between PDS, non-PDS and missing^a
Missing	16 (28%)	13 (22%)	11 (19%)	17 (30%)	
Non-PDS	24 (24%)	30 (31%)	33 (34%)	11 (11%)	$\chi^2 = 14.14$; $p = .028$
PDS	24 (27%)	31 (35%)	16 (18%)	18 (20%)	
Gender	Male(n, %)	Female (n, %)			
Missing	43 (34.4)	82 (65.6)			
Non-PDS	134 (47.9)	146 (52.1)			$\chi^2 = 7.03$; $p = .030$
PDS	72 (47.4)	79 (52.6)			
Type of education	ULO (n, %)	HBO (n, %)	Different... (n, %)	Quitted/graduated (n, %)	
missing	31 (66%)	14 (30%)	0	2 (4%)	
Non-PDS	32 (43%)	30 (40%)	5 (67%)	8 (10%)	$\chi^2 = 17.36$; $p = .008$
PDS	25 (32%)	35 (45%)	3 (3%)	14 (18%)	

^a Pearson Chi-Square ^b within non-PDS versus PDS and missing

Statistical analyses

Missing values analyses

Previous research (Maulana et al., 2017) revealed that there are significant differences gender differences in behavior of pre-service teachers. If the missingness in our data points towards a gender imbalance, this could influenced the results in this study. The proportion of missing cases (individuals) for each questionnaire at each time point was examined by univariate missing values analyses using SPSS. At time point 1, the proportion of missing values was around 50-70%; teacher efficacy ($n = 441$; 65.5% missing), stress ($n = 363$; 53.9% missing), student questionnaire 'My teacher...' ($n = 344$; 55.1% missing) and peer observation ($n = 478$; 71% missing). At time point 2, the proportion of missing values was around 80-90%; teacher efficacy ($n = 559$; 83.1%), stress ($n = 543$; 80.7%), student questionnaire 'My teacher...' ($n = 527$; 78.3%) and peer observation ($n = 609$; 90.5%). At time point 3, the proportion of missing values was around 88-93%; teacher efficacy ($n = 594$; 88.3%), stress ($n = 583$; 86.6%), student questionnaire 'My teacher...' ($n = 571$; 84.8%) and peer observation ($n = 629$; 93.5%).

Differences in the proportion of missing values between non-PDS-based teachers and PDS-based teachers were examined, and it appeared that the percentage of missing values for some instruments were higher among the PDS teachers. This was found for teacher efficacy, stress and

student questionnaire 'My teacher...' at time point 1, 2 and 3. For the peer observations, the proportion of missing values did not differ between PDS and non-PDS teachers.

There appeared to be no differences in the proportion of missing values between men and women on the main outcome variables included in the study. In addition, differences between age groups on the proportion of missing values were not found. However, for some of the variables at time point 1, the younger respondents seemed to be less likely to have missing values, compared to older respondents. For time points 2 and 3, no such trend was found. Contrary to the findings on time point 1, the younger respondents were more likely to be missing in the data set at time point 3.

Reliability analyses

Scale reliabilities (Cronbach's alpha, Item-scale correlations) were computed for the subscales of the student questionnaire and the peer observation. The reliabilities were sufficient to high except for the support scale in the survey on job satisfaction.

See Appendix A.

Descriptive statistics

Descriptive statistics (mean, standard deviations) were computed for PDS and non-PDS teachers separately, at time point 1, time point 2 and time point 3. ANOVA's were used to compare the groups.

Effect size computations were used to indicate the substantive importance of the differences between PDS and non-PDS pre-service teachers. Effect sizes were calculated using Hedge's g (g). An effect size of 0.20 was considered to be a small effect, 0.50 a medium, and 0.80 a large effect, according to Cohen's rule of thumb. A positive effect size indicates that the value is higher for the PDS pre-service teachers, while a negative sign indicates a higher value for the non-PDS pre-service teachers.

Results

Descriptive statistics

Student perceptions of their teachers' teaching skills

The average teaching skills of PDS teachers were rated significantly higher by their students, compared to the teaching skills of non-PDS teachers (Table 2). At time point 1 the effect size was 0.50 but remained not significant (most likely due to a statistical power problem caused by the relatively small sample size). These differences were significant at time point 2 with a medium effect size of 0.56. At time point 3, the effect size had changed in the advantage of the non-PDS teachers, but again this was not significant.

In the longitudinal study, the student perceptions of pre-service teachers in the non-PD-schools increased in the first year from 1.79 to 2.1 (effect size of 0.23). The scores of the pre-service teachers in the PD-schools increased between time point 1 and 2 also about a quarter of standard deviation from 2.32 to 2.56 (effect size 0.22).

Table 2 Warm's-theta score differences between PDS and non-PDS teachers on at teacher level of the student questionnaire *My teacher...*

	Time point 1						Time point 2						Time point 3					
	<i>m</i> non- PDS (<i>n</i> =15 9)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =80)	<i>sd</i>	<i>P</i> ^a	<i>g</i> ^b	<i>m</i> non- PDS (<i>n</i> =53)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =18)	<i>sd</i>	<i>p</i>	<i>g</i>	<i>m</i> non- PDS (<i>n</i> =53)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =23)	<i>sd</i>	<i>p</i>	<i>g</i>
Warm's Theta Student questionnaire (cross-sectional)	1,57	1,24	1,71	1,2	0,13	0,11	2	0,92	2,48	0,83	0,08	0,54	2,15	0,89	2,08	0,96	0,5	-0,07
Warm's Theta Student questionnaire (longitudinal)	1,79	1,02	2,32	1,14	0,13	0,5	2,1	0,82	2,56	0,79	0,05	0,56	2,23	0,9	2,07	0,91	0,08	-0,17

^a *P*-value = significance of the independent samples t-test ^b *g* = effect size (Hedge's *g*): .20 = *small*, .50 = *medium*, .80 = *large*

Peer observation of teachers' teaching skills

Results of the peer observations were generally in line with that of the student perceptions. Scores of PDS teachers were significantly higher compared to that of non-PDS teachers and at time point 2 the differences were significant with a medium to large effect size (Table 3a).

In the longitudinal study, the pre-service teachers trained in PDS-trajectories, who stay at the school of their internship, revealed higher scores (effect size = 0.49), but the difference was not significant (due to a statistical power problem caused by the relatively small sample size). One year later the beginning teachers in PDSs showed significantly higher scores compared to the teachers in non-PDSs. This effect was not observed in the third year.

The scores of pre-service teachers in the non-PDSs grew from .62 to 1.28 (effect size of .38), while the scores of the pre-service teachers in the PDSs increased between time point 1 and 2 with more than a full standard deviation from 1.39 to 2.66.

To ensure that these results were not influenced by the attraction and retention of the better teachers, by PDSs compared to non-PDSs, we did the following analysis: the mean scores of the teachers who remained in the school were compared in measurement 1. This revealed that the starting skill levels of the teachers who remained, were equal for teachers in the PDS and the non-PDS (see second row of Table 3a).

Additionally, a multi-level analysis (using SPSS) was conducted to reveal the magnitude of influences at the level of the school and the teacher. The results are summarized in Table 3b. Model 1 reveals that males seem to show more increase in their teaching quality compared to females, and the teaching quality of females was scored slightly higher by the observers (effect size: 18; n.s.). Model 2: teaching skills during the teacher training and gender together explain about 26% of the growth on teacher level after one year and around 68% of the growth on school level. Model 3: being trained in a PD-school, together with gender and skill as a student teacher, explains approximately a quarter of the growth on teacher level and nearly 100% of the growth on school level.

Table 3a Differences between non-PDS and PDS teachers on ICALT peer observation

	Time point 1						Time point 2						Time point 3					
	<i>m</i> non-PDS (<i>n</i> =145 ^c)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =38 ^d)	<i>sd</i>	<i>p</i> ^a	<i>g</i> ^b	<i>m</i> non-PDS (<i>n</i> =42)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =21)	<i>sd</i>	<i>p</i>	<i>g</i>	<i>m</i> non-PDS (<i>n</i> =33)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =10)	<i>sd</i>	<i>p</i>	<i>g</i>
Warm's Theta Observation (cross-sectional)	.93	2.00	1.18	1.56	.54	.13	1.43	1.71	2.53	1.83	.03	.63	1.35	1.55	1.98	1.25	.25	.43
Warm's Theta Observation (longitudinal)	.62	1.70	1.39	1.16	.15	.49	1.28	1.84	2.66	1.29	.02	.80	1.55	2.14	1.52	1.57	.98	-.02

^a*P*-value = significance of the independent samples t-test ^b*g* = effect size (Hedge's *g*): .20 = *small*, .50 = *medium*, .80 = *large* ^c sample size fluctuates per subscale, between 78 and 90 for non-PDS and 55 and 60 for PDS.

Table 3b Explanatory factors of growth in teaching skills after one year

		model 0			model 1			model 2			model 3		
-2* log likelihood		253.20			249.65			156.65			153.00		
		est	se	sig									
unexplained variance	Teacher	2.54	.53	.00	2.39	.50	.00	1.87	.47	.00	1.90	.40	.00
	School	.68	.48	.16	.70	.46	.15	.22	.32	.50	.00	.00	
intercept: teaching skill Θ		1.77	.28	.00	1.78	.27	.00	1.17	.27	.00	1.25	.24	.00
female (1)					-.40	.21	.06	-.58	.22	.01	-.61	.21	.01
teaching skill as a student teacher Θ								.55	.14	.00	.56	.13	.00
professional development trajectory											.49	.23	.04

Teacher stress

Only one of the stress causes showed significant differences: PDS teachers perceived significantly more learning opportunities at time point 1, compared to non-PDS teachers ($p = 0.03$; $g = -0.27^2$) (see Table 4). The stress outcomes were also similar in both groups.

² The negative sign is due to the item formulations, in the direction of lacking opportunities

Table 4 Differences between non-PDS and PDS teachers on the teacher stress questionnaire

	Time point 1						Time point 2						Time point 3					
	<i>m</i> non-PDS (<i>n</i> =164)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =93)	<i>sd</i>	<i>p</i> ^c	<i>g</i> ^d	<i>m</i> non-PDS (<i>n</i> =85)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =35)	<i>sd</i>	<i>p</i>	<i>g</i>	<i>m</i> non-PDS (<i>n</i> =60)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =24)	<i>sd</i>	<i>p</i>	<i>g</i>
High psychological task demands	1.43	.29	1.48	.30	.23	.17	1.49	.29	1.44	.28	.37	-.17	1.55	.32	1.66	.31	.15	.35
Lack of learning opportunities	.75	.38	.65	.36	.03	-.27	.82	.43	.66	.40	.06	-.38	.83	.46	.69	.48	.19	-.30
Lack of regulating possibilities	1.47	.34	1.44	.33	.44	-.09	1.47	.32	1.47	.29	.96	.00	1.46	.35	1.51	.40	.57	.14
Poor social-organizational aspects	.9	.27	.87	.28	.34	-.11	.88	.28	.94	.22	.26	.23	.88	.30	.97	.39	.26	.27
Tension	.39	.24	.40	.22	.89	.04	.35	.23	.43	.23	.08	.35	.38	.22	.49	.27	.07	.47
Discontent	.25	.21	.30	.23	.41	.23	.24	.19	.28	.21	.31	.20	.25	.21	.30	.23	.41	.23
Job satisfaction	1.32	.42	1.45	.55	.23	.28	1.38	.52	1.43	.46	.65	.10	1.32	.42	1.45	.55	.23	.28

^a*P*-value = significance of the independent samples t-test ^b*g* =effect size (Hedge's *g*): .20 = *small*, .50 = *medium*, .80 = *large*

Teacher efficacy

Differences between PDS and non-PDS teachers were evident for some of the teacher efficacy scales. PDS teachers perceived significantly higher teacher efficacy regarding instructional strategies ($p = 0.02$; $g = 0.25$) and regarding engaging students in learning ($p = 0.01$; $g = 0.37$) compared to non-PDS teachers at time point 1 (see Table 5). At time point 2, PDS-based teachers perceived higher levels of collective teacher efficacy regarding instruction, compared to non-PDS teachers. At time point 3, differences were found in favor of the PDS teachers for three teacher efficacy scales, respectively, instruction, class management, and student engagement, with large effect sizes (see Table 5 for more details).

Table 5 Differences between PDS and non-PDS teachers on the teacher efficacy questionnaire

	Time point 1						Time point 2						Time point 3					
	<i>m</i> non-PDS (<i>n</i> =128)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =71)	<i>sd</i>	<i>p</i> ^c	<i>g</i> ^d	<i>m</i> non-PDS (<i>n</i> =73)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =34)	<i>sd</i>	<i>p</i>	<i>g</i>	<i>m</i> non-PDS (<i>n</i> =59)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =19)	<i>sd</i>	<i>p</i>	<i>g</i>
Efficacy for instructional strategies	3.34	.41	3.48	.38	.02	.35	3.43	.39	3.45	.30	.79	.05	3.44	.33	3.73	.45	.00	.80
Efficacy for classroom management	3.19	.55	3.37	.70	.07	.30	3.39	.54	3.44	.50	.67	.09	3.44	.54	3.77	.66	.03	.58
Efficacy for student engagement	3.19	.43	3.35	.45	.01	.37	3.30	.35	3.25	.30	.53	-.15	3.25	.39	3.50	.52	.03	.59
Collective efficacy for instruction	3.43	.49	3.50	.52	.35	.14	3.32	.52	3.50	.35	.04	.38	3.29	.47	3.50	.50	.11	.44
Collective efficacy for classroom management	3.56	.47	3.64	.53	.23	.16	3.42	.47	3.60	.40	.05	.40	3.46	.49	3.53	.55	.64	.14

^a*P*-value = significance of the independent samples t-test ^b *g* = effect size (Hedge's *g*): .20 = *small*, .50 = *medium*, .80 = *large*

Job satisfaction

At all time-points no differences were found between PDS-based teachers and non-PDS-based teachers with regard to job satisfaction (see Table 6).

Table 6 Differences between PDS and non-PDS teachers regarding job satisfaction

	Time point 1						Time point 2						Time point 3					
	<i>m</i> non-PDS (<i>n</i> =120)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =67)	<i>sd</i>	<i>p</i> ^c	<i>g</i> ^d	<i>m</i> non-PDS (<i>n</i> =48)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =27)	<i>sd</i>	<i>p</i> ^c	<i>g</i> ^d	<i>m</i> non-PDS (<i>n</i> =15)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =20)	<i>sd</i>	<i>p</i> ^c	<i>g</i> ^d
Job conditions	3.06	.74	3.08	.68	.86	.03	3.00	.73	2.77	.66	.19	-.32	3.27	.74	2.93	.82	.21	-.43
Support	3.56	.53	3.61	.57	.53	.10	3.41	.50	3.42	.53	.98	.01	3.34	.49	3.41	.55	.69	.14
Autonomy	3.94	.56	3.88	.70	.55	-.09	3.91	.47	3.91	.56	.98	-.01	3.94	.63	3.84	.83	.67	-.15
Relations	3.90	.60	3.82	.67	.43	-.12	3.82	.60	3.85	.63	.81	.06	3.89	.69	3.94	.65	.82	.08
Work	3.89	.67	3.91	.81	.89	.02	3.90	.74	4.06	.75	.39	.21	4.05	.89	3.82	1.00	.47	-.25

^a*P*-value = significance of the independent samples t-test ^b*g* = effect size (Hedge's *g*): .20 = *small*, .50 = *medium*, .80 = *large*

Job Commitment

At all time-points no differences were found between PDS-based teachers and non-PDS-based teachers with regard to job-commitment (see Table 7).

Table 7 Differences between PDS and non-PDS teachers regarding job-commitment

	Time point 1						Time point 2						Time point 3					
	<i>m non-PDS (n=117)</i>	<i>sd</i>	<i>m PDS (n=63)</i>	<i>sd</i>	<i>p^c</i>	<i>g^d</i>	<i>m non-PDS (n=48)</i>	<i>sd</i>	<i>m PDS (n=27)</i>	<i>sd</i>	<i>p</i>	<i>g</i>	<i>m non-PDS (n=15)</i>	<i>sd</i>	<i>m PDS (n=20)</i>	<i>sd</i>	<i>p</i>	<i>g</i>
Continuance commitment	3.30	1.28	3.13	1.33	.41	-.13	3.79	1.40	3.54	1.32	.44	-.19	3.09	1.28	3.90	.99	.41	.72
Normative commitment	3.12	1.22	3.14	1.20	.90	.02	3.08	1.07	3.00	1.28	.78	-.07	3.39	1.50	3.4	1.26	.90	-
Affective commitment	5.65	.81	5.67	.97	.87	.03	5.53	1.02	5.86	.92	.17	.34	5.78	.95	5.37	1.08	.87	.40

^aP-value = significance of the independent samples t-test ^bg = effect size (Hedge's g): .20 = small, .50 = medium, .80 = large

Satisfaction with teacher preparation

PDS-based teachers were significantly more satisfied with regard to the teacher preparation they received compared to non-PDS-based teachers. PDS-based teachers rate the content of the program, acquired practical skills, instructions, the assessment, study schedule, and institute contribution to the program significantly more positive compared to non-PDS-based pre-service teachers (see Table 8).

Table 7 Differences between PDS and non-PDS teachers regarding teacher preparation

			Time point 1		<i>p</i> ^a	<i>g</i> ^b
	<i>m</i> non-PDS (<i>n</i> =97)	<i>sd</i>	<i>m</i> PDS (<i>n</i> =65)	<i>sd</i>		
Judgement content education	3.12	.74	3.35	.65	.04	.33
Judgement acquired general skills	3.38	.81	3.53	.74	.21	.20
Judgement acquired academic skills	2.91	.78	2.91	.81	1.00	.00
Judgement acquired practical skills	3.05	.82	3.42	.83	.01	.45
Judgement career preparation	3.64	.78	3.70	.85	.64	.07
Judgement instructors education	3.37	.80	3.69	.69	.01	.42
Satisfaction information education	3.03	.87	3.22	.88	.17	.22
Satisfaction study facilities	3.44	.62	3.57	.56	.18	.22
Satisfaction assessment and evaluation education	3.17	.79	3.62	.73	.00	.58
Satisfaction study schedule	3.45	.82	3.72	.86	.05	.32
Satisfaction study load	3.00	.82	3.02	.93	.90	.02
Satisfaction study support/guidance	3.28	.91	3.52	.80	.09	.28
Satisfaction involvement institute	2.77	.77	3.05	.85	.03	.35

^a*P*-value = significance of the independent samples t-test ^b*g* = effect size (Hedge's *g*): .20 = *small*, .50 = *medium*, .80 = *large*

Conclusion and discussion

PDSs in the Netherlands are partnerships that aim to: 1) bridge the gap between the job requirements and the theoretical curriculum requirements and 2) to reduce the theory/practice gap experienced by novices. This is to be attained by close collaboration between the teacher education institutes and the internship schools. The aim of this study was to evaluate the behavioral and the psychological outcomes of teacher preparation based in PDS in secondary education in the Netherlands over a period of 3 years.

Our findings confirm hypothesis 1, showing that the quality of teaching skills of PDS teachers were evaluated more favorably by their students, compared to the teaching skills of non-PDS teachers in the second year of the study only. Peer observations were in line with the findings with regard to the perceptions of students; experts evaluated the teaching skills of PDS teachers higher compared to non-PDS teachers in the second year only. This is in line with the findings of Castle et al. (2006), Sharpe et al. (1999), Houston et al. (1999), and Ridley et al. (2005). The decline in the rating of the teaching skills measured at time point 3 may indicate reduced support in PDSs after the first year on the job. During this data collection period, most schools in the Netherlands supported their beginning teachers during the first year after certification, leading to a permanent contract or to exit. After a permanent contract is achieved, the school support comes to an end or reduces dramatically. The exception to the rule is when schools supply comprehensive induction programs. At the time of this data collection (ending in 2012), comprehensive induction programs were scarce in the Netherlands.³ We therefore speculate that the lack of support in the third year is reflected in the outcome of this study. The decline in teaching skills at time point 3 could also be due to the low response rate (longitudinal morbidity).

Our study confirms hypothesis 2, revealing that differences were found between PDS and non-PDS teachers on teacher efficacy. PDS teachers generally reported more positive levels of teacher efficacy. This was shown in the first, second, and third year of the study. These results confirm the results reported by Sandholtz et al. (2000), and Swars and Dooley (2010). Our finding suggests that PDS programs seem to be beneficial for preservice teachers' self-efficacy. Science teachers research indicates that preservice teachers graduating with higher levels of self-efficacy tend to make better

³ The need for ongoing support after certification has been revealed in many international studies referred to induction arrangements for beginning teachers. In this regard the partnership between schools and teacher education institutes becomes continuous.

progress as beginning teachers (Appleton & Kindt, 2002). Hence, we argue that PDS programs should be maintained and improved further in the future.

With regard to research question 3, no differences were found in stress causes and stress responses between PDS and non-PDS teachers, with the exception of the amount of learning opportunities, which was perceived as less by non-PDS teachers compared to PDS teachers in the first data collection year (the teacher preparation year). It seems that PDS-based pre-service teachers experience more learning opportunities compared to non-PDS-based pre-service teachers. The lack of impact of PDS on the other stress causes and stress responses might be attributed to equal support offered in this regard in both tracks. A rival explanation is that more support and the focus on development and accountability might add to perceived job demands. The combined effect of more resources as well as higher job demands might result in a zero net effect on the perceived stress causes and responses. Future research needs to disentangle this issue as it is speculative at this point in time.

No differences were found between PDS-based and non-PDS-based teachers with regard to job satisfaction and job-commitment (research questions 4 and 5). The same explanation as for the lack of influence on stress might hold for job satisfaction and commitment too. Future research might shed more light on this issue.

The collaboration between the teacher education institutes and the schools is associated with more positive evaluations of the teacher preparation program compared to non-PDS-based teachers (research question 6). This means that pre-service teachers are more satisfied with their preparation for the job rating the content of the program, acquired practical skills, instructions, the assessment, study schedule and institute contribution to the program, significantly more positive compared to non-PDS-based pre-service teachers. These results confirm the research that showed that teachers who are trained well and who are certified, perform better than teachers who are not certified and are only trained in the setting of working as teachers (Boyd et al., 2009; Darling-Hammond & Youngs, 2002; Dutch Educational Inspectorate, 2011; Dronkers, 2010).

The behavioral and some of the psychological outcomes point to the advantage of teachers following internships in PDSs in the Netherlands. The teachers develop their teaching skills more rapidly but this stagnates after two years. The third year stagnation should be studied more closely in the future by comparing PDS teacher preparation with and without comprehensive induction arrangements.

We realize that our study cannot make causality claims due to the design used. However, the evidence provided in this study gives indications favoring PDS trajectories with regard to effective

behavior development and a smoother transition to the teaching profession. This study reveals positive associations only. All associations point in the direction that pre-service teachers benefit from being educated and training in the context of a PDS: a context where the teacher education institute and schools work closely together and each has its added value. Future investments in the co-operation between schools and teacher education institutes seem warranted.

References

- Appleton, K., & Kindt, I. (2002). Beginning elementary teachers' development as teachers of science. *Journal of Science Education*, *21*, 155–168.
- Barton, L., Barrett, E., Whitty, G., Miles, S., & Furlong, J. (1994). Teacher education and teacher professionalism in England: Some emerging issues. *British Journal of Sociology of Education*, *15*(4), 529-543.
- Blocker, L. S., & Mantle-Bromley, C. (1997). PDS versus campus preparation: Through the eyes of the students. *The Teacher Educator*, *33*, 70-89.
- Book, C. L. (1996). Professional development schools. In J. Silula, T. Buttery, & E. Buyton (Eds.), *Handbook of research on teacher education* (pp. 194-210). New York: Macmillan.
- Boyd, D. J., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis*, *31*(4), 416–440.
- Canrinus, E. T. (2011). *Teachers' sense of their professional identity*. Enschede, The Netherlands: Groningen: PrintPartners Ipskamp B.V.
- Canrinus, E. T., Helms-Lorenz, M., Beijaard, D., Buitink, J., & Hofman, A. (2012). Self-efficacy, job satisfaction, motivation and commitment: exploring the relationships between indicators of teachers' professional identity. *European Journal of Psychology of Education*, *27*(1), 115–132.
<http://dx.doi.org/10.1007/s10212-011-0069-2>.
- Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, *44*(6), 473–490. doi:10.1016/j.jsp.2006.09.001
- Castle, S., Fox, R., & Souder, K. (2006). Do professional development schools (PDSs) make a difference? A comparative study of PDS and non-PDS teacher candidates. *Journal of Teacher Education*. Retrieved from <http://jte.sagepub.com/content/57/1/65.short>
- CFI (2005). Subsidieregeling dieptepilot voor de opleidingsschool en de academische school 2005-2008. OCW regeling Algemeen Verbindend voorschrift.
- Cooper-Hakim, A., & Viswesvaran, C. (2005). The construct of work commitment: Testing an integrative framework. *Psychological Bulletin*, *131*, 241-259.
- Corcoran, E. (1981). Transition shock: The beginning teacher's paradox. *Journal of Teacher Education*, *32*, 19–23.

- Creemers, B., & Kyriakides, L. (2008). The dynamics of educational effectiveness. Retrieved from http://samples.sainsburysebooks.co.uk/9781134168392_sample_520130.pdf
- Darling-Hammond, L. (2005). The design of teacher education programs. ... *Teachers for a* Retrieved from <http://news.cehd.umn.edu/wp-content/uploads/2009/06/PreparingTeachersforaChangingWorldChpt11.pdf>
- Darling-Hammond, L., & Youngs, P. (2002). Defining “highly qualified teachers”: What does “scientifically-based research” actually tell us? *Educational Researcher*, 31(9), 13–25.
- Davis, J., & Wilson, S. (2000). Principals’ efforts to empower teachers: Effects on teacher motivation and job satisfaction and stress. *The Clearing House*. Retrieved from <http://www.tandfonline.com/doi/pdf/10.1080/00098650009599442>
- den Brok, P. J., Wubbels, T., & van Tartwijk, J. W. F. (2017). Exploring beginning teachers' attrition in the Netherlands. *Teachers and Teaching*, DOI:10.1080/13540602.2017.1360859
- Dronkers, J. (2010) De relatie tussen leerkrachtentekort en de taal- en natuurkundekennis en –vaardigheden van 15-jarige leerlingen. Maastricht: ROA.
- Duffin, L. C., French, B. F., & Patrick, H. (2012). The teachers' sense of efficacy scale: Confirming the factor structure with beginning pre-service teachers. *Teaching and Teacher Education*, 28, 827-834. doi: 10.1016/j.tate.2012.03.004
- Dutch Ministry of Education, Culture and Science (2002). *De school centraal [The school in the center]*. The Hague, the Netherlands: Ministry of Education, Culture and Science.
- Dutch Ministry of Education, Culture and Science (2008). *Landelijke criteria Opleiden in de School [National criteria for Educatin within Schools]*. KPMG group.
- Educational Inspectorate (2007). *Opleiden in de school. Kwaliteitsborging en toezicht [Education within schools. Quality assurance and control]*. Rijswijk: Den Haag media groep.
- Evers, A. (2013). Individueel assessment. In W. Shaufeli, & A. Bakker (Eds.), *De psychologie van arbeid en gezondheid* (pp.91-110). Houten: Bohn Stafleu van Loghum.
- Flores, M. (2001). Person and context in becoming a new teacher. *Journal of Education for Teaching: International Research and Pedagogy*, 27(2), 135-148. Retrieved from <http://www.tandfonline.com/doi/pdf/10.1080/02607470120067882>

- Fuller, F. F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal*, 6(2), 207–226.
- Hargreaves, A. (2000). Four ages of professionalism and professional learning. *Teachers and Teaching: Theory and Practice*, 6(2), 151-182. doi:10.1080/713698714
- Harmsen, R., Maulana, R., Helms-Lorenz, M., & van Veen, K. (2017). *Beginning teachers' perceived stress: Causes, responses, and relationships with teaching behavior and attrition*. Paper presented at AERA 2017.
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Retrieved from <http://books.google.nl/books?hl=nl&lr=&id=vSoUT6PXdoIC&oi=fnd&pg=PP1&dq=Hattie+2012&ots=XITb-quaPU&sig=E8HjGdb0hFI9K4oK6VUFsu2sCTE>
- Helms-Lorenz, M., van de Grift, W., & Maulana, R. (2015). Longitudinal effects of induction on teaching skills and attrition rates of beginning teachers. *School Effectiveness and School Improvement*, 27(2), 1–27. <http://dx.doi.org/10.1080/09243453.2015.1035731>.
- Houston, W. R., Hollis, L. Y., Clay, D. Ligons, C., & Roff, L. (1999). Effects of collaboration on urban teacher education programs and professional development schools. In D. Byrd & D. J. McIntyre (Eds.), *Research on professional development schools: Vol. 7. Teacher education* (pp. 6-28). Thousand Oaks, CA: Corwin Press.
- Huberman, M. (1989). The professional life cycle of teachers. *Teachers' College Record*, 91, 31–57.
- Ingersoll, R. M., & Strong, M. (2011). The impact of induction and mentoring programs for beginning teachers: A critical review of the research. *Review of Educational Research*, 81, 201-233. doi: 10.3102/0034654311403323
- Kruger, T., Davies, A., & Eckersley, B. (2009). Effective and sustainable university-school partnerships. *Beyond determined efforts by inspired individuals*. Retrieved from http://www.aitsl.edu.au/docs/default-source/aitsl-research/insights/re10022_effective_and_sustainable_university_school_partnerships_beyond_determined_efforts_by_inspired_individuals_vic_uni_mar_2009.pdf
- Kyriakides, L., Creemers, B. P. M., & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education*, 25(1), 12–23. doi:10.1016/j.tate.2008.06.001
- Lambert, R. G., McCarthy, C., O'Donnell, M., & Wang, C. (2009). Measuring elementary teacher stress and coping in the classroom: Validity evidence for the classroom appraisal of resources and demands. *Psychology In The Schools*, 46(10), 973-988. doi:10.1002/pits.20438

- Landsman, M. J. (2001). Commitment in public child welfare. *Social Service Review*, 3, 386-419.
- Latham, N. I., & Vogt, W. P. (2007). Do professional development schools reduce teacher attrition?: Evidence from a longitudinal study of 1,000 graduates. *Journal of Teacher Education*, 58(2), 153-167.
- Mantle-Bromley, C. (2001, April). *The status of early theories of professional development school potential: Progress of four settings across the United States*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
- Mantle-Bromley, C., Gould, L. M., McWhorter, B. A., & Whaley, D. C. (2000). The effect of program structure on new teachers' employment and program satisfaction patterns. *Action in Teacher Education*, 22(1), 1-14.
- Maulana, R., & Helms-Lorenz, M. (2016). Observations and student perceptions of the quality of preservice teachers' teaching behaviour: Construct representation and predictive quality. *Learning Environments Research*, 19(3), 335-357. <http://dx.doi.org/10.1007/s10984-016-9215-8>.
- Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2014). Development and evaluation of a questionnaire measuring teaching behaviour: A Rasch modelling approach. *School Effectiveness and School Improvement*, 26(2), 169-194 Advance online publication.
- Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2015). A longitudinal study of induction on the acceleration of growth in teaching quality of beginning teachers through the eyes of their students. *Teaching and Teacher Education*, 51, 225-245. <http://dx.doi.org/10.1016/j.tate.2015.07.003>.
- Maulana, R., Lorenz, M., & van de Grift, W. (2017). Validating a model of effective teaching behaviour of pre-service teachers, *Teachers and Teaching. Theory and Practice*, 23(4), 471-493. <http://dx.doi.org/10.1080/13540602.2016.1211102>.
- Meyer, J. P., Allen, N. J., & Smith, C. (1993). Commitment to organizations and occupations: Extension and test of a three-component conceptualization. *Journal of Applied Psychology*, 78, 538-551.
- Nath, J. L., Guadarrama, I. N., & Ramsy, J. (2011). *Investigating University-school Partnerships*. IAP. Retrieved from <https://books.google.com/books?id=Ve96jJUsUs0C&pgis=1>
- National Council for the Accreditation of Teacher Education. (2001). *Standards for professional development schools*. Washington. Retrieved from <http://www.ncate.org/ProfessionalDevelopmentSchools/tabid/497/Default.aspx>.
- NVAO (2009). *Toetsingskader opleidingsschool* [Assessment framework professional development schools]. The Hague, the Netherlands: NVAO.

- Onderwijsraad (2001). *Educatieve infrastructuur: Onderzoek in opdracht van de onderwijsraad*. Den Haag: IVA Tilburg/Onderwijsraad.
- Paese, P. C. (2003). Impact of professional development schools pre-service through induction. *Action in Teacher Education*, 25(1), 83-88. doi:10.1080/01626620.2003.10463296
- Patterson, J. H. (2000, November). *Impact of professional development schools on teacher education*. Paper presented at the annual meeting of the Mid-South Educational Research Association, Bowling Green, KY.
- Ploeg, J. D. van der, & Scholte, E. M. (2003). *Arbeidssatisfactie in het onderwijs en de jeugdzorg*. Houten: Bohn Stafleu Van Loghum
- Ridley, D. S., Hurwitz, S., Davis Hackett, M. R., & Knutson Miller, K. (2005). Comparing PDS and campus-based pre-service teacher preparation: Is PDS-based preparation really better? *Journal of Teacher Education*, 56(1), 46-56.
- Ross, F. (2001, April). *So what type of teachers are they? Graduates of a PDS teacher preparation program 3-6 years later*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
- Sandholtz, J., & Dadlez, S. (2000). Professional Development School Trade-Offs in Teacher Preparation and Renewal. *Teacher Education Quarterly*, 27(1), 7-27. Retrieved from <http://www.jstor.org/stable/23478205>
- Sharpe, T. L., Lounsbery, M., Golden, C., & Deibler, C. (1999). Analysis of one ongoing district-wide collaborative approach to teacher education. *Journal of Teaching in Physical Education*, 19, 79-96.
- Swars, S., & Dooley, C. (2010). Changes in teaching efficacy during a professional development school-based science methods course. *School Science and Mathematics*. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1949-8594.2010.00022.x/full>
- Thompson, R., & Ross, F. (2000). Becoming a teacher in a professional development school. *Teaching and Change*, 8, 31-50.
- Timmermans, M. (2012). *Kwaliteit van de opleidingsschool. Over Affordance, Agency en Competentieontwikkeling*. Retrieved from <http://www.onderwijsdatabank.nl/79498/>
- Tournaki, N. (2009). Pathways to teacher certification: Does it really matter when it comes to efficacy and effectiveness? *Action in Teacher Education*, 30(4), 96-109. Retrieved from <http://www.tandfonline.com/doi/pdf/10.1080/01626620.2009.10734455>

- Trachtman, R. (1996). *The NCATE professional development school of study: A survey of 28 PDS sites*. Unpublished manuscript. (Available from Professional Development School Standards Project, National Council for Accreditation of Teacher Education, Washington, CD 20036).
- Tschannen-Moran, M. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, 68(2), 202-248. Retrieved from <http://rer.sagepub.com/content/68/2/202.short>
- Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing and elusive construct. *Teaching and Teacher Education*, 17, 783-805.
- Tusin, L. (1995). *Success in the first year of teaching: Effects of a clinical experience program*. Paper presented at the annual meeting of the Association of Teacher Educators, Detroit, MI.
- van de Grift, W. (2007). Quality of Teaching in Four European Countries: A review of the literature and an application of an assessment instrument. *Educational Research*, 49(2), 127–152.
- van de Grift, W., van der Wal, M., Torenbeek, M., (2011) Development in teaching skills, 88(6), 416–432. Retrieved from http://apps.isiknowledge.com/full_record.do?product=UA&search_mode=GeneralSearch&qid=1&SID=T1d73nDzVf33dWXJX8e&page=1&doc=10.
- van de Grift, W., Helms-Lorenz, M., & Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Studies in Educational Evaluation*, 43, 150–159.
- Veldhoven, M., van (1996). *Psychosociale arbeidsbelasting en werkstress*. [Psychosocial job demands and job stress]. Lisse: Swets & Zeitlinger.
- Veldhoven, M., van, Jonge, J. de, Broersen, S., Kompier, M., & Meijman, T. (2002). Specific relationships between psychosocial job conditions and job-related stress: A three-level analytic approach. *Work & Stress*, 16(3), 207-228.
- Veldhoven, M., van & Meijman, T.F. (1994). Het meten van psychosociale arbeidsbelasting met een vragenlijst: De vragenlijst beleving en beoordeling van de arbeid (VBBA) [The measurement of psychosocial job demands with a questionnaire: The Questionnaire on the experience and evaluation of work (QEEW)]. Amsterdam: NIA.
- Wait, D. B. (2000). *Are professional development school trained teachers really better?* Paper presented at the annual meeting of the National PDS Conference, Columbia, SC.

Wait, D. B., & Warren, L. L. (2001). *Are professional development school trained teachers better classroom managers?* East Lansing, MI: National Center for Research on Teacher Learning.

Walling, B., & Lewis, M. (2000). Development of professional identity among professional development school preservice teachers: Longitudinal and comparative analysis. *Action in Teacher Education*, 22(2), 65-72.

Yerian, S., & Grossman, P. L. (1997). Preservice teacher perceptions of their middle level teacher education experience: A comparison of a traditional and a PDS model. *Teacher Education Quarterly*, 24(4), 85-101.

Appendix A: Scale descriptives

Questionnaire /observation	Scales	Number of items	Reliability	N
Commitment	Continuance commitment	6	.72	100
	Normative commitment	6	.69	112
	Affective commitment	6	.69	108
Job satisfaction	Fringe-benefits and salary	6	.79	135
	Support on the job	8	.50	133
	Amount of job-autonomy	6	.80	132
	Relationships at work	5	.80	134
	Tasks	5	.77	133
ICALT student questionnaire	Providing a safe and stimulating learning environment	12	.90	2064
	Activating learning	17	.95	2013
	Classroom management	8	.89	2108
	Teaching learning strategies	10	.85	2050
	Adaptive teaching	6	.84	2106
	Clarity of instruction	11	.71	2029
	Subject knowledge and skills	4	.74	2126
	Student involvement	10	.86	2114

	Effort	64	.87	2123
	Social desirability	68	.63	2136
	All variables teaching quality InVO		.98	1711
ICALT observation	Safe and stimulating learning climate	4	.80	120
	Efficient organization of the lesson	4	.81	120
	Clear and structured instruction	7	.88	115
	Intensive and activating lesson	7	.80	109
	Adapting instructions to relevant differences	4	.71	102
	Teaching of learning strategies	6	.89	100
	Involvement of students	3	.87	120
	All variables	35	.95	86
NSE	Satisfaction with quality of the program	9	.85	111
	Satisfaction with acquired skills	6	.84	112
	Satisfaction with acquired scientific skills	7	.92	25
	Satisfaction with acquired practical skills	5	.89	111
	Satisfaction with career preparation	5	.86	111
	Satisfaction with quality of lectures/educators	9	.91	111
	Satisfaction with study facilities	9	.85	111
	Satisfaction with assessment procedures	4	.75	111
	Satisfaction with the schedule	3	.86	111
	Rating of the study load	3	.77	111
	Satisfaction with the provided support	2	.81	108
	Satisfaction with the personal engagement with the			

	institute	4	.88	110
Perceived stress	High psychological task demands	32	.85	143
	Lack of learning opportunities	10	.82	145
	Lack of regulating possibilities	23	.83	140
	Poor social-organizational aspects	45	.91	137
	Tension	51	.93	129
	Job discontent	21	.81	125
Self-efficacy	Classroom self-efficacy	19	.90	106
	School self-efficacy	14	.83	109
	TMWH instruction self-efficacy	8	.75	108
	TMWH classroom management self-efficacy	8	.92	106
	TMWH student engagement self-efficacy	9	.84	104
	Collective self-efficacy instruction	6	.87	104
	Collective self-efficacy discipline	6	.89	104